•

Met Ala Val Asp Val Arg Ile Ala Ala Phe Leu> GACCAATCGGAGTGTGAATTTTTGGAAAATATTTTGT<u>Q</u>CGGTTCCTTTAGTTGTAATATACTACTTTACAAATTTGGAATAATTTGGCAGCAAAAACCATCTGCAGCAACAA AATCATCTGCAGCTGCGAAATCATCTGCAGCAGCAAAAAGCATCTTCAGGAGCGAGAAAAAGCCCCAAATAATGTGAG ATG GCA GTT GAC GTC CGA ATC GCT GCC TTC CTG CCT GCA GTT TTA GCT CAA GAG AGA TGT GGG TAT ATG ACC GCC ATC CCA AGG CTA CCA CGA CCG GAT AAT TTG CCA Pro Ala Val Leu Ala Gln Glu Arg Cys Gly Tyr Met Thr Ala Ile Pro Arg Leu Pro Arg Pro Asp Asn Leu Pro> ATA ACA GCC AAC CTC GGC ACG CAG GTC ATC TAC ATG GAT GAA GAG ATA GAA GAC GAA ATC ACC ATC GCC ATA CTT AAT TAT AAC GGA CCA Ile Thr Ala Asn Leu Gly Thr Gln Val Ile Tyr Met Asp Glu Glu Ile Glu Asp Glu Ile Thr Ile Ala Ile Leu Asn Tyr Asn Gly Pro> GTA CTA AAT TIT GAA GGC CAG ACA TGG AGT CAG AGG CCC CTG CTC CCC GCC GAG CGG GAT GAC CTG TGC ATG GAC GCC TAC CAC GTG Val Leu Asn Phe Glu Gly Gln Thr Trp Ser Gln Arg Pro Leu Leu Pro Ala Pro Glu Arg Asp Asp Leu Cys Met Asp Ala Tyr His Val> TCA ACT CCG TTC ATT GAA CTG CCA TTT TTA TCC GGT TCG TAC AAT CTG CTG ATG CCG GTC ATC AGG AGA GTT GAC AAC GGG AGT GCA TCT Ser Thr Pro Phe Ile Glu Leu Pro Phe Leu Ser Gly Ser Tyr Asn Leu Leu Met Pro Val Ile Arg Arg Val Asp Asn Gly Ser Ala Ser> CAT CAT CAC GCA AGA CAG CAT TAC GAG TTG CCC GGC ATG CAG CAG TAC ATG TTC AAT GTG CGC GTG GAC GGC CAG TCG CTG GTG GCA GGC His His Ala Arg Gln His Tyr Glu Leu Pro Gly Met Gln Gln Tyr Met Phe Asn Val Arg Val Asp Gly Gln Ser Leu Val Ala Gly> GTG TCT CTC GCT ATC GTC AAC ATA G.T GAC AAC GCG CCC ATC ATA CAA AAC TTC GAG CCT TGC CGG GTT CCT GAA CTG GGC GAG CCA Ser Leu Ala Ile Val Asn Ile Asp Asp Asn Ala Pro Ile Ile Gln Asn Phe Glu Pro Cys Arg Val Pro Glu Leu Gly Glu Pro TTG ACA GAA TGC ACA TAC CAA GTA TCG GAC GCG GAC GCA CGC ATC AGC ACA GAG TTC ATG ACG TTC AGG ATC GAC AGC GTT CGT GGC Leu Thr Glu Cys Thr Tyr Gln Val Ser Asp Ala Asp Gly Arg Ile Ser Thr Glu Phe Met Thr Phe Arg Ile Asp Ser Val Arg Gly THE RESERVE OF THE PARTY OF THE Leu Val Phe Ile Ala Pro Ala Val Leu Ala Gln TTT ATA GCG CTG GTG

Asn Thr Ser Leu Asn Phe>

6 6

GAG GAG ACC TTC TAC ATC GAA CGG ACG AAT ATC CCC AAC CAA TGG ATG TGG CTA AAT ATG ACC ATA GGC GTT AAT ACC TCG CTC AAC 'TTC Glu Glu Thr Phe Tyr Ile Glu Arg Thr Asn Ile Pro Asn Gln Trp Met Trp Leu Asn Met Thr Ile Gly Val Asn Thr Ser Leu Asn Phe

<u>.</u> 4		. A		ے ۵		~ ^		۸		۸ م		. ^		^		^		. ^			
AAT		GCG ATC Ala Ile>		GIY		AAA Lys>		GAA Glu>		GCT Ala>		ACC Thr>		AAC Asn>		CAG Gln>		CAT His		CAT His>	
GCG				CCT		TAC		AAG Lys		A.A.C A.s.n		CAG Gln		GAC Asp		GTC Val		GAG Glu		TAC Tyr	
GTG Val		AGG Arg		CTC Leu		GCT		CAC His		CAA Gln		CGA Arg		ACC		GCC		GTC		AAC 'Asn	
CAA Gln	1120	crc crc val	1210	GCC Ala	1300	GTC	1390		1480	GGT Gly	570		1660	GCG /	50	CAC (His A	40		30	TTC A Phe A	50
GTG	Н	ACA	7	GAG Glu	ਜ	ATC (Ţ	CCT ATA Pro Ile	14	TTA (Leu (15	TAC C	16	GTA G	17	GAA C	184	GAT GTC Asp Val	193	TCC T Ser P	2020
ATG (TTC Phe				Acg 1		GAA G		GAT 1 Asp 1		GGC T Gly T		GTG G Val V		TTC G		GGG G		GAC TV Asp So	
ATG A	0		0 (AGC ATT Ser Lie	0	CTT A	0 +	E⊢o	0 •	C) o	٥.	GTC G Val G	0.		0 *	ATC T Ile P					_
ACT A	1110	caa aac Gln asn	1200	TTC A Phe S	1290	CCA C Pro L	1380	AGA CARG	1470	GAT A Asp L	1560	GAA G	1650	ATT CGG Ile Arg	1740	CCG A'	1830	GAC ATC Asp Ile	1920	GCT AAC Ala Asn	2010
GTG A Val T		TAC C		TTC T		TTT C															
										T CAT e His		G CCT a Pro		T ACG e Thr		s cAG		c AGA		TCA Ser	
	1100		1190	c ACA P Thr	1280		1370	c GAC	1460	GGA TTT Gly Phe	1550	A GCG	1640		1730	r GAG	1820	GAC Asp	10	GTC	2000
c cAC r His	1	3 AAA u Lys		A GAC	. .	A GAG J Glu	7	AAC Asn	à		7	ATA	ñ	AGT Ser	Ħ	GAT	18	CAC	191	CGC Arg	20
c Acc		A GAG		GAA Glu		CGA Arg		ATC Ile		TTC		TAC		cAG Gln		AAC Asn		GCA Ala		ATC Ile	
AA(GAA		GAG		CAA Gln		GAC Asp		GAA Glu		TTC		TTT Phe		TGG Trp		GTT Val		GAC	
CCC Pr	1090	TTT	1180	ACA AAT Thr Asn	1270	CTG	1360	ACA Thr	1450	AAA Lys	1540	GCA Ala	1630	GAG Glu	1720	AAT Asn	1810	GCG	1900	ссс 61у	066
TGG CTG CCG. AAC 'SET 'Leu' PIU "ASh		CAG Gln				ACA Thr	, ,	GTG Val		GAT Asp	.,	GAG		CCA		ATC Ile	7	AAG Lys	-	ACC Thr	7
TCG Ser		CAA Gln		ATC Ile		GAC Asp		ATT Ile		TTC Phe		GCT Ala		GTG Val		CTC		GCC		CTC	
GAC	080	GTC	1170	CTC	1260	CGC Arg	\$0	ATC Ile	1440	AAC Asn	1530	GCT	50 *	ÀΒ	*		0° *		0 *		086
Eeu Eeu	10	GCT	11	AGG Arg	12	GAC	13	GTC Val	14	CTC	15	GGC Gly	1620	TAC	1710	ATT GAC Ile Asp	1800	TTC GTC Phe Val	1890	GAC AAA Asp Lys	198
ACA GCC.		TTC Phe		TAC Tyr		ATT		GTC		ACC		CCA Pro		GAT Asp		CAC H		TTC 7		ATC C	
ACA Thr	0 *	ATC Ile	0 *	AAC Asn	0 *	CCA Pro	0 *	AAC (Asn 1	0 *	CTG 7		CCT (TC en		GTT C		βG 1.γ		ACC A	
3TG. /81	107		116		125(TCG (1340	ACA 1	1430	CCC (1520	GAC (Asp E	1610	ATG C Met I	1700	TTG C	1790	cAG G	1880	CTG A	1970
TTC AGC GTG. Phe Ser Val		CTG GAG Lau Glu		CCT ATC Pro Ile		GTG 1 Val S		TCA A		ACG C		0 A		TCC A		GCG TAIL		осс с с1у с		TTC C'Phe L	
TTC A				TG C		TC G eu V		CA T		GAG A Glu T		CGT CTA GAG AGC G'G Arg Leu Glu Ser Va.								TTG GGT AAC GCT GTT AAC TTC Leu Gly Asn Ala Val Asn Phe	
ra T le P	09	SC T	¥ 20	AT A sn M	Q *	ည် ဂျ	o *	SC A	0 *	o o n	o *	NG A(o *	E E	0 *		<u>o</u> *	9 G	0 *	T AS	0
AGT CCG CTG CAT ATA Ser Pro Leu His Ile	1060	cce cgc TGG Pro Arg Trp	1150	ATC AAT ATG Ile Asn Met	1240	GCT GTA TTC CTC Ala Val Phe Leu	1330	TTC TCC ACA Phe Ser Thr	1420	ATC ATG GAG Ile Met Glu	1510	6 G	1600	CTC AAT CAC Leu Asn His	1690	CAC GTG GGC GTC His Val Gly Val	1780	GAG ACT GAA Glu Thr Glu	1870	TTG GGT AAC GCT GTT AAC Leu Gly Asn Ala Val Asn	1960
S P H C						F &		S e E E		C AT		1 C1		r CI		C GT s Va		წნ ა		c 60	
c cTC o Leu		T CCG g Pro		GAC ACT GAG Asp Thr Glu				G GCC						Acc Thr				GAC ASP		AA(
AGT CCG Ser Pro	1050	AGC CGT Ser Arg	1140	GAC ACT Asp Thr	1230	AGC GGG Ser Gly	1320	GAG GAG Glu Glu	1410	CTG GCA Leu Ala	1500	ACG GTG Thr Val	1590	ATG GGC Met Gly	1680	ACG AGG Thr Arg	1770	ACC TTC Thr Phe	1860	651 G1y	1950
Sel:	, ,		,,	GA(GAC Glu		CTC	-		-	ATC ATG GGC Ile Met Gly	-	GAC ACG AGG Asp Thr Arg	7		-	TTG	-
ACC Thr		AAC Asn		GGA Gly		AAA Lys		GAT		CGA		CAG TAC Gln Tyr	•			GAC		GTC Val		TTA	
GTC Val	±040	GTG Val	1130	GAC Asp	1220	GGA G1y	1310 *	TAT Tyr	₹ 400 *	TAT Tyr	490	CAG Gln	1580	TTC	1670	AAC Asn	1760	ACC	1850	ACT	1940

.

				<i>ې</i> ز
^ ^		Λ Λ		cc caa la clus ac aca sp args
CGA CTA Arg Leu> - ATC ACC Ile Thr>	GGC CGC Gly Arg> CGC GTT Arg Val>	AAC ACC Asn Thr> ACT CTG Thr Leu>	TAC AAC TYR ASDA AAC ACA ASD THE	GCG GAA Ala Glu> GAC AGA Asp Arg.
			CTC TAC AAC Leu Tyr Asn GTG AAC ACA VAL ASN ThE GAT CCT GCA	CAG GCG GIN Ala
c ATA 1 11e 0 c GTC s Val	c GGC	CTT CTT CTT CTT CTT CTT CTT	CCG CTC Pro Leu 2650 TCC GTG Ser Val 2740 GAA GAT	CAG CTT Leu
CAG CTG GTC Gln Leu Val 2110 GAT QGC CAC ASP QLY HIS	2200 AGC ACC AAG CAA Ala Thr Lys Gln 2290 TG GCT ATC GGC Leu Ala Ile Gly	CTG CGT GAC CTT AAC ACC Val Arg Asp Leu Asn Thr: 2470 TCG GTG GAG GGG ACT CTG TTP Val Glu Gly Thr Leux	2560 * GAC GGA CCG CTC ASP Gly Pro Leu 2650 * CAA ATT TCC GTG Qln Ile Ser Val 2740 * TCG ACA GAA GAT * * * * * * * * * * * * * * * * * * *	2830 GTC CCG CAG Val Pro Gln 2920 AGT GAT CTT Ser Asp Leu
CAG CTG Gln Leu 2 GAT GGC ASP G1Y	C AC a Th c GC		C GGA P GLY A ATT B LIG	G GTC S Val
TCA CF Ser G100		SG GTG FIGURE 10 A T T T T T T T T T T T T T T T T T T	FC GAC B ASP TO GAC	in Lys
GCG ACG TCA CAG Ala Thr Ser Gln 2100 AAC GTG CCT GAT ASN VAI Pro Asp	ACC TCT TTCM GCC ACC AAG Thr Ser Phe Ala Thr Lys 2280 2290 AAC GG GGA CTG GCT ATC ASD ANG GLY Leu Ala Ile	2370 GTG AGG GTG Val Arg Val 2460 CCG GTG TGG Pro Val Trp	2550 GAC GAC ATC Asp Asp Ile 2640 CTC ACG GGT Leu The Gly 2730 2730 Ann Ann Ann Cych	CAC ACG AAC AAG GTC CCG CAG GCG GAA ASP Thr Asn Asn Lys Val Pro Gln Ala Glu 2910 2920 2910 2920 GTC ACT CTG ATA GCC AGT GAT CTT GAC AGA GTC ACT Ile Ala Ser Asp Leu Asp Arg.
GCG ACG Ala Thr 21 AAC GTG ASN VAl	ACC TThr S AAC C ASD A		Sp Ac CC	ACG AG Thr Ag ACT CT Thr Left
thr P	SAC ASP TASP TASP TASP TASP ASP ASP ASP ASP ASP ASP ASP ASP ASP	TTC A Leu T	40 GCG G Ala A 30 * TTC C Phe L	10 GAC A ASP T ASP T GTC A GTC A
CAC ACG GCG ACG TCA CAG His Thr Ala Thr Ser Gln -2090 . 2100 GAG GAG AAC GTG CCT GAT Glu Glu Asn Val Pro Asp	2180 TGG GAC Trp Asp 2270 ATT AAC	2360 TCC CTC ACA Ser Leu Thr 2450 GAC AAC GCG	2540 2550 2560	2810 2820 ACC GAC ACG AAC AAC Thr Asp Thr Asn Asn 2900 2910 GTC GTC ACT CTG ATA Val Val Thr Leu Ile
		CTC Leu Leu AAC AAC Asn Asn	GTG over Nat Arc carc carc carc carc carc carc carc	ATC / Ille 1
### CCC TTC Glu Pro Phe	ATA Ile CCC Pro	GTC Val ATG Met	TCC GTG Ser Val ATG ATC Met Ile GTG GTC Val Val	CAC ATC His Ile GAC GAG Asp Glu
GAG ANGGERE GOG GAA TCCC TTC ASP Thr Leu Gly Glu Pro Phe 2070 2080 CCT CGA GGC AGT CCC CAA GTG Pro Arg Gly Ser Pro Gln Val	2170 TTC GAG ATA Phe Glu Ile 2260 ATC TTC CCC Ile Phe Pro	2340 TAC GAG GAG TTT GAG GTC CTC TCC CTC ACA TYR Glu Glu Phe Glu Val Leu Ser Leu Thr 2430 2440 ATA ACT ATA ATC GAT ATG AAC GAC AAC GCG Ile Thr Ile Ile Asp Met Asn Asp Asn Ala	2530 2540 2550 2560 GTG GTG GGC GCG GAC GAC ATC GAC GCA CCG CTC TAC AAC Val Val Gly Ser Val Arg Ala Asp Asp Ile Asp Gly Pro Leu Tyr Asn 2620 2630 2640 2650 GAC CTG ATA ATG ATC GAC TTC CTC ACG GGT CAA ATT TCC GTG AAC AACA ASp Leu Ile Met Ile Asp Rhe Leu Thr Gly Gln Ile Ser Val Asp Thr Tac TAC ACG GGT CAA ATT TCC GTG AACA ATT TCC TACA AACA A	ATC ACA ATC CAC ATC ACC Ile Thr Ile His Ile Thr 2890 ACC CAC TTA GAC GAG GTG Thr His Leu Asp Glu Val
FTG CGG, Leu Gly 2 GGC AGT Gly Ser	TTC Phe ATC Ile	7TT TTT Phe 2 2 ATC Ile	GTG Val	ACA Thr CAC CAC His
Leu Leu GGC	GCC Arg Acc	GAG Glu Glu		ATC ILE
GAG ACKG ASP Thr 2070 CCT CGA Pro Arg	2160 GAT CTG Asp Leu 2250 ATC GAG Ile Glu	2340 TAC GAG TYr Glu 2430 ATA ACT	2520 GGG CTC GLY Leu 2610 GAT AAG ASP LYS 2700 2700 CAC CTC His Leu	2790 2800 GGA AAT ATC ACA ATC Gly Asn Ile Thr Ile 2880 2890 AAC GCA ACC CAC TTA Asn Ala Thr His Leu
* Ala Thr Ala Thr 50 * CGG CTG Arg Leu	G GCC rr Ala rg GAA	A GAC e Asp C ACA u Thr	GCG GGC Ala Gly CACA Asp Thr Asp Thr CGC TTC CGC TTC	
EGA GET ACR Arg Ala Thr 2060 TTA CGC CTG Leu Arg Leu	2150 ACC ACG GCC Thr Thr Ala 2240 TGC GTG GAA Cys Val Glu	2330 ACC ATA GAC Thr Ile Asp 2420 ATG CTC ACA Met Leu Thr	2510 TCG GCG GGC Ser Ala Gly 2600 GAA GAC ACA Glu Asp Thr 2690 CCA CGC TTC	2780 2780 GAA ACC Glu Thr 2870 ATT TAC Ile Tyr
GTG CONTRACT THE LE			ATG TC 2 CGT GP GILL SE GILL S	2780 TGG GAA ACC Trp Glu Thr 2870 TAT ATT TAC Tyr Ile Tyr
CAG GTG CGA AIG Thr GIN Val Arg AIG Thr 2060 CCC ACC TTA CGC CTG Pro Thr Leu Arg Leu	CCC GAC Pro Asp AGG AAT Arg Asn	AAC GTG Asn Val GAA TCG	GAG ATG TCG GCG GGC Glu Aet Ser Ala Gly 2600 CCT CGT GAA GAC ACA Pro Arg Glu Asp Thr 2690 ACT CCT CCA CGC TTC Thr Pro Pro Pro Arg Phoe	TAT TY
2050 CCA C	2140 ACC GAC CCC Thr Asp Pro 2230 CAG TTT AGG Glu Phe Arg		2500 CGC G Arg G 2590 TTC C Phe P 2680	2770 * ACT T Thr T 2860 : GTC G
TTT GTG Phe Val 2050 ACG CCA Thr Pro	2140 ACC GAC Thr Asp 2230 GAG TTT Glu Phe	2320 AGA CAC Arg His 2410 TAC GAC	CGA GTC CGC Arg Val Arg 2590 ACC ATT TTC Thr Ile Phe 2680 GAC GCG QAT	2770 CCG ACT Pro Thr 2860 ACC GTC
TTA Leu AAC	GCC Ala	ATC 1 Ile 1 GAC 3 ASP 5	Arg Arg Thr Thr Thr	GAC GAS GAS GAT ASP ASP ASP ASP
AGT GAA TTA TTT GTG Ser Glu Leu Phe Val 2040 2050 ATC AAC AAC ACG CCA Ile Asn Asn Thr Pro	2130 2140 TTA CGC GCC ACC GAC CCC Leu Arg Ala Thr Asp Pro 2220 2230 AAC CCC GAC GAC TTT AGG Asn Pro Asp Glu Phe Arg	2310 c GAA g Glu 2400 A GAC	2490 in Phe in Phe 2580 2580 2580 2570 2670	2760 2770 CCC CCT GAC CCG ACT Pro Pro Asp Pro Thr 2850 2860 AAG TTC GAT ACC GTC Lys Phe Asp Thr Val
		S A S C C C C C C C C C C C C C C C C C	2490 CAG AAC TTC Gln Asn Phe 2580 GTG CGA TAC Val Arg Tyr 2670 2670	2760 2770 TGC CCC CCT GAC CCG ACT Cys Pro Pro Asp Pro Thr 2850 2860 ACT AAG TTC GAT ACC GTC Thr Lys Phe Asp Thr Val
CGA GAA Arg Glu 0 * AAT GAC Asn Asp	GAG Glu GCT Ala	GCG Ala TAC TYr		0 GAT TGC ASP CYS 0 ACG ACT Thr Thr
CGA Arg Arg 2030 AAT Asn	2120 CAG Gln 2210 CAG	2300 GTA Val 2390 GTC	2480 GAG GBU 2570 CAA GIN 2660 AGC	2750 GAT Asp 2:840 ACG ACG
				•

•	GGC Gly>	CTC Leu>	GCT Ala>	ccG Pro>	CTG Leu>	∢ ô	υ ຈ	, G &	F. 0	- A	`
			AAT GCT Asn Alae			T ATA E Ile>	G TTC u Phe>	C TTG e Leu>	c TTT r Phe>	r cgc	
	G ACC		ic AZ	c GCC e Ala	G GAG o Glu	A <u>C</u> AT e H18	r GAG o Glu	3 TTC u Phe	3 TAC 3 Tyr	3 ATT	
0	A GAG	0 * C TTC e Phe	0 * T GAC n Asp		0 G CCG 1 Pro	0 * F ATA a Ile	0 * G CCT A Pro	o A GAG Y Glu) A CGG 1 Arg	A ACG	2 -
3010	AAC CGA Asn Arg	3100 * ATC TTC Ile Phe	3190 * GTG AAT Val Asn	3280 * CAT ATC His Ile	3370 GAA GTG Glu Val	3460 TAC GCT TYr Ala	3550 * TAC GCG Tyr Ala	3640 AAC GGA Asn Gly	3730 * TCA CAA Ser Gln	3820 * CGG ATA Arg Ile	3910
			3 T GTC P Val								
_	c crc	c cgr	G GAT u Asp	A CCA u Pro	C ATC p Ile	G ACG Y Thr	c TAC n Tyr	A GTG r Val	G GAA u Glu	3 TTC 1 Phe	
3000	C TCC	3090 * ACG CAC Thr His	3180 c TTC u Leu	3270 * CTT GAA Leu Glu	3360 CGG GAC Arg Asp	3450 TGG GGG Trp Gly	3540 * TTC AAC Phe Asn	3630 * GCG ACA Ala Thr	3720 * GAT GAG ASP Glu	3810 * G GAG g Glu	3900
	c TTC e Phe		31 C TTG e Leu				o TT			AG Ar	
	c TTC n Phe	A CCA u Pro	T ATC 1 Ile	c cGT 1 Arg	G GAG r Glu	A TAT Y TYF	r ccg	r cra	4 GGC 1 Gly	3 ATC 1 Ile	
2990	G AAC t Asn	3080 * GAT GAA Asp Glu	3170 trc GTT eu Val	3260 * GC GTC	3350 * GC ACG er Thr	3440 * AAG GGA Lys Gly	3530 * ATC CAT Ile His	3620 * GGA GTT Gly Val	3710 * GTC GTA Val Val	3800 * AG GAG lu Glu	3890
7	G ATG u Met		0 1	9 9	A S					9 9	38
	A CTG	r ggr p gly	A GTT 1 Val	s cAG	CTC Leu	CTC CTC	ATC Ile	AAT Asn	CAA Gln	CCA	
[]	r cga	o T GAT 9 ASP	o A GAA C Glu	r AAG	S AAC	GAC	CTG	ATC Ile	TTC	GTT Val	
	AAC CCT Asn Pro	3070 * GAC CGT ASP Arg	3160 * GAC ACA Asp Thr	3250 * AAC CTT Asn Leu	3340 ATC CTG Ile Leu	3430 * GCC ATG Ala Met	3520 * TAT GAG Tyr Glu	3610 * GCT GTA Ala Val	3700 GTC ACC Val Thr	3790 CAA GCC Gln Ala	3880
1, I											
(Q EN	A GTG a Val	cTG	3 AAC	r GAG	GAG	ACC Thr	ACA Thr	CGA Arg	GTC	CTG	
1	TAT GCA Tyr Ala	3060 * GAG GTG Glu Val	3150 * AAT CAG Asn Gln	3240 * 'A TCT 'e Ser	3330 * GGC TAC Gly Tyr	3420 * *C GAG *U Glu	3510 * AAC GAG Asn Glu	3600 * G GAA g Glu	3690 * G GGC a Gly	3780 * KG TTA G Leu	3870
ļ.				A1 I1		ភូ ភ្ន		AG Ar	A1	AG Ar	m
i ii	AAT Asn	r 660 6 61y	AGA Arg	ACT Thr	GTC Val	GAG Glu	ATG . Met	, GCG	CAC	TTC	
109	ATC Ile)50 * r AGT	.40 * r AAC	3230 TCT TGG	3320 * CC AGG	10 * GGA G1Y	3500 * TG TCC let Ser	3590 * GA CTT rg Leu	3680 * 3GA CTC	70 * TCG Ser	*
. C.	Tac Gre Tar Val	30g GGT 1 Gly	314 GGT G1Y	F 03	332 TCC	341 ACG	350 CAA ATG Cln Met	U Æ,	368 GAC GGA Asp Gly	377 GGC Gly	386
	AGC TAC GTC Ser Tyr Val	CAG	GAA Glu	CTC	GAC AAC Asp Asn	GTC Val	CAA Gln		GAC	CTC	
		GAG ACC	3130 * GGG GAA GGA Gly Glu Gly	GAA	3310 * CCC GAC ACA GAC AAC TCC Pro Asp Thr Asp Asn Ser	3400 CAG ATC GCG AAC Gln Ile Ala Asn	3490 * GGC ATT CCG Gly Ile Pro	GTC Val	3670 * GCG ACT GAT CCG GAC GGA Ala Thr Asp Pro Asp Gly	3760 * GGC GAG AAC Gly Glu Asn	
2950	CAC A <u>c</u> G GTG His Thr Val	3040 GAC TAT GAG ASP TYR Glu	3130 ccc caa Gly Glu	1 3220 ccg AGC Pro Ser	3310 * CCC GAC ACA Pro Asp Thr	3400 ccc Ala	3490 * ATT	3580 * GAT GCC Asp Ala	3670 * CAT ASP	3760 * GAG Glu	3850
	A CO	TAT TYE		CCC Pro	GAC	ATC Ile	66C 61y	GAT	ACT		
	CAC His		ATG	CCG			CAC His	AAC Asn	GCG	GAT	
2940	GAA ATA TAC Glu Ile Tyr	3030 TAC GTG TYr Val	3120 GAC AAC TTC ATG ASP ASN Phe Met	3210 TTG CCA Leu Pro	3300 * CGC GAC GAG Arg Asp Glu	3390 * ATC ATA Met Ile	3480 * TTC GAC CAC Phe Asp His	3570 3580 * * CCG ACC AAC GAT GCC GTC Pro Thr Asn Asp Ala Val	3660 * CGG ATA TCG Arg Ile Ser	3750 * GTT AAC GAT Val Asn Asp	3840
~	ATA Ile		AAC Asn	3 TTG Leu	3 GAC Asp	3. ATG Met	3. TTC Phe		3. ATA Ile		e.
		GTC Val		GAA		G'IG Val	o 3480 3490 351 * * * * * * * * * * * * * * * * * * *	TTC Phe		0 * CAA GTA Gln Val	
2930	GAC Asp	3020 * CTG Leu	3110 * ATC Ile	00 * CCT Pro	90 GAC ASP	80 TTT Phe	70 CgC	60 TC GTC Val	50 cAG Glu	40 * CAA Gln	30
5.9		30	31	3200 * Ç	3290 * G	3380 *	3470 *	3560 * G	3650 * G	3740 * CA G1	3830

٠, ٠	•		•	,		-						J									
TTC Phe>		AAC Asn>	AAC Asn>		ATT Ile>		GGG Gly>		TAC Tyr>		ATC Ile>		Grc val>		ACC Thr>		ACG Thr>	4830	AACG	0764	
AGA Arg		AAG Lys	CGC		66c 61y		GCA Ala		GAC Asp		AAT Asn		ACC Thr		GAC Asp		AGC		AAAC		
CCT		ATC Ile	GTT Val		GGT Gly		ACC Thr		ATA Ile		CTT Leu		GTC Val		GCG		CAC His	4820	ATAC	40	
GAA Glu	4000	GAC	4090 * CCT Pro	4180	GAT Asp	4270	TAC Tyr	4360	GCT Ala	4450		4540		4630	ATC Ile	4720	GAG Glu	8	TCG	4940	
GGA Gly	4	CAA GAC Gln Asp	4090 * GAT CCT ASP Pro	4	CCC GAT	4	TTG TAC Leu Tyr	4	TAT GCT Tyr Ala	4	CTG ACG Leu Thr	4	ACC GAC Thr Asp	4	TTT	4	CTC C		rgag		
CAA		GCA			TCG		GAA		ACT		GTC Val		CGC Arg		GAC ASP		GCG (4810	A ACAGATCCGTAGTGACCTAGTCCTCCTGAGCTCGATACAAACAA	4930	
ACG Thr	06	CTA Leu	4080 C GGC CTG e Glylleu	4170	AAC Asn	4260	AGG Arg	4350	ATT Ile	4440	GGA G1y	4530		4620	AGA Arg	4710	GTG Val		AGTC		
CCC	39	CCT Pro	40 TTC Phe	41	AGT AAC Ser Asn	42	GTG Val	43	GCC	44	ACC Thr	45	CAG GAC Gln Asp	46	AAC AGA Asn Arg	47	GGC	00 *	ACCT	20	
GTG Val		CTT	CAT His		GCT		TTT Phe		TCG		CAA Gln		GCT		GAC		ACC	4800	AGTG	4920	
TTT	0 *	CAA Gln		0 *	cc 1a	o *	GTC Val	40		0 *		0 *	GGC Gly	0 *	GAA	0 *			CCGT		
GTT Val	3980	CAC His	4070 * GAA GGT Glu Gly	4160	GTG Val	4250	CCA Pro	434	GAA GGC Glu Gly	4430	AAC GCT Asn Ala	4520	GCC	4610	GTC	4700	CCC GTC Pro Val	4790	AGAT	4910	
GTT Val		TCT Ser	AGC		CAA Gln		CGT Arg		TCT		CTC		ACG		CAG		GAC Asp		A AC, XXX>	•	
AGA Arg		GAG Glu	AAC Asn		CTG		CCT		CAG Gln		GTA		GAC Asp		CAA G		AAC (Asn 1		TAG /	00	
TTC: AGA 	3970	GAA		4150	_	4240	GAC	4330	ACC	4420		4510	ACT	4600	CTG (4690	GCT A	4780		4900	
Ac@ Thr	m	ATG	4060 * GAT GGC Asp Gly	4	CAC	4	GCA	4	GCG	4	GCT TTC Ala Phe	4	GCT	4	ACG Thr 1	4	CCC (4	ATG AGA Met Arg		
A Te		GGC Gly	ATC Ile		TCC		GAG Glu		CAT		TCG (Ser		ACA (Thr)		AAC A		GTG (CTG ATG Leu Met	4890	
GAC ATG ACG ASP Met Thr	096	GCC A		0 *	GCC	0 *	AGG (0.*	TTA Leu	0 *	cAG Glu	0 *	GTC /	o *	GTC /	<u>o</u> *		o *	TCG (Ser I	4	
Thr	39	AGT (4050 * CGT ATT Arg Ile	4140	AGT (4230	GTG Z	4320	AGA '	4410	AGA (4500	GAA (4590	TTC (4680	CAA GTG Gln Val	4770	TAC 1	0	
Ser		AAG	TAT		CAA		ACC (Thr		CTC		GTG 1		TTC (Phe (GTC 1		GAC (Asp (CCG 1 Pro 1	4880	
CTG TCC Leu Ser	0 *	GAA G	0 * TAC '	0 *	AA 1n	0 *	GTT 1	0 *	CTG (0 *	GCA (0 *	AAA 1 Lys I	0 *	TTC (Phe V		Ic le	-	TAC (Tyr F		
	395(ATA (Ile (404(ATT Ile	413		422(ACT (Thr)	431(4400	GNG (Gla 1	449(4580	TAC 1	4670		4760		4870	
3GA		mrc /	AGC 1		TA /		TC /		AGA GAG Arg Glu		CTC C		CTG TTC Leu Phe		GTC 1		TGC AAC Cys Asn		ACA ACG Thr Thr	4	
CCA (GCT GTC GCT TTC Ala Val Ala Fie			TTC CTG AAG AAA GAG CTG ATA AGG Phe Leu Lys Llys Glu Leu Ile Arg		CT (000 1 61y 1		AGC (Ser I		GGA G		cgc c		ACC 1			0	
AC (3940	TC (4030 TGT (4120	AG C	4210	eu 1	4300		4390	CC P	4480	AT G	4570	AC C	4660	TG A	4750	AC G	4860	
rce c	36	GCT GTC Ala Val	40 13AC 7	4	A S	42	TC C	43	TCC ATC Ser Ile	43	AC C	4	TG C	4.	CAG AAC CGC Gln Asn Arg	46	AAC ATG Asn Met	47	TCA TAC GGG Ser Tyr Gly		
GA 4		CAT (His A	4020 4030 * * TGT GAA GAC GAC TGT CAC Cys Glu Asp Asp Cys His		AG A		CC A		GAC TCC ATC Asp Ser Ile		GTG GAC CCC AGC Val Asp Pro Ser		ACG GCC ACG ATG CAT Thr Ala Thr Met His		TCG C		TTC A		ACT TCA TAC Thr Ser Tyr	4850	
AG G	0 *	AA O	o PAP G	0 *	TG A	0 *	CT T	0 *	CG G	0 *		0 *		0 *		0 1	GG T	0 +		4	
Sp Gs	3930	TCA GAA CAT Ser Glu His	4020 cr GA ys Glu	4110	TC C	4200	CT G	4290	ACA GCG GAC Thr Ala Asp	4380	ATG GTA Met Val	4470	CG G	4560	GTA TCC Val Ser	4650	GCT GGG Ala Gly	4740	GCG GCC Ala Ala	0	
GCT ACA GAC CAG GGA ACG GAC CCA GGA CCG ALLERA					TTC T		CCA CIT CCT GCT TCC ATC CIT ACT GTC Pro Leu Pro Ala Ser Ile Leu Thr Val				ACA A'		CCC ACG GCC Pro Thr Ala				TTC AGC GCT GGC TTC AAC ATG ACC Phe Ser Ala Gly Phe Asn Met Thr			4840	
GCT ACA Ala Thr		GCG TCC Ala Ser			AGG T		CA C		ATA TCC Ile Ser		GAT AC ASP T		CAG CC		TAC GTG Tyr Val		TTC AC		cag Arg Gln Met		
ΰα	3920	७ द	4010 * C2 H3	4100	A A	4190	ΣĞ	4.280	¥Ή	4:370	¥ G	4460	ฮ ซิ	4550	8 6	4640	. F. C.	4730	: 0 G		
				-		-		-		•		-		-		•		•			

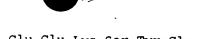
										5570	2560
NTTATT	TAATTTAAGA	ATACTCGATT	AGTCATTAAG	AAATGCCATT	aaattttata	TCATCTCATC	SCTCTTGCTC	CTCCATTTAA	AAATTGAGAC	ATTGAAAAAA	TTTAAAACTGTAATTGAAAAAAAAAATTGAG ACCTCCATTTAAGCTCTTTTGCTCTCATCTCA
5550	5540	5530	5520	5510	5500	5490	5480	5470	5460	5450	5440
TCTCATT	gtcataag <u>c</u>	AAGGATAGTT	AGATTTTTGA	TTTGGAATTA	AATTTTGGAG	TAAGATCATT	CAACACTAGT	TAGAAGATAA (CAGAAAGATC	AACGGACAGTT	cctgagticgttaacggacagticagaaagatctagaagataacaacactagttaagatcattaatttttggagtittggaattaagatttttgaaaggatagttgtgataag <u>cc</u> tgtgatt
5430	5420	5410	5400	5390	5380	*	*	*	*	*	*
CTTCTC	ACCCGACTC	TTCG <u>C</u> TTTCA	CAACAACAAC	TCGCAAACCA		5370	5360	5350	5340	5330	5320
5310	5300	1 5290	5280	×	TATCCCGAAG	GGCAAATGGT 5370	TGAACGAAAA 5360	SATCCTGAGG 5350	CGT CTTCATG	ACTTTGGCAA	GAAGATCTTG COCCACTTTGCCAACGT CTTCATGGATCCTGAGGTCAAAGGCAAATGGTTATCCCGAAGTCGCAACCACAACAACAACTTCGCTTTCAACCCGACTCCCTTTCTCG 5320 5330 5340 5350 5360 5370 5380 5390 5400 5410 5420 5430
PCGGCATC	CTGATCTGA	TCCAACGACT	199909999	5270	5260 * TATCCCGAAG	5250 * GGCAAATGGT	5240 TGAACGAAAA 5360	5230 GATCCTGAGG	5220 CCTCTTCATG 5340	5210 gactificgcaa	5200 * AAGATCTTG <u>COX</u> 5320
.5190	5180			TAGATGCCAT	ACGCCAGAIT 5260 TATCCCGAAG	AGCAATAAAG 5250 GGCAAATGGT 5370	TCTTCAATGA 5240 TGAACGAAAA 5360	TCCAACCCTA 5230 \$ATCCTGAGG	TGTGGAAGGC 5220 cczcttcatg	LACAAACACAC 5210 \$ACTTTGGCAA 5330 i	SCCCGGCACC 5200 * AAGATCTTG COM
SCATCGC	Accececed	5170	5160 *	5150 * TAGATGCCAT	5140 ACGCCAGATT 5260 TATCCCGAAG	5130 AGCAATAAAG 5250 GGCAAATGGT	5120 * TCTTCAATGA 5240 * TGAACGAAAA	5110 TCCAACCCTA 5230 SATCCTGAGG	5100 TGTGGAAGGC 5220 * CGTCTTCATG	5090 AACAAACACAC 5210 ACTITGCAA	5080 5090 5110 5120 5130 5140 5150 5160 5170 5180 5190 GCCCCCGGCACCAACAACAACACCTATCTTCAATGAAGCAATAAAGACCAATAAAGACCAATTAGATGCATTAGCGAGGTTCCAACGACTCTGATCTTCATCGATCTTCAATCGATTATCCCAATTTAGATTCTTCAACCTAAAAGCAAATGCTTATCCCAAAGTCGAAATCGTATTCCCAAACACAACAACAACAACAACAACAACAACAAC
5070		TCTGGATTGA	CTCACTGGAC 5160 *	CCAAG <u>TAC</u> GG 5150 TAGATGCCAT 5270	CTGTCGATGA 5140 ACGCCAGATT 5260 TATCCCGAAG	GCTAAACCGACGGTTGGAAGCCCTGTCGATGACGAAGTACGCTCACTGGACTCGATTGAACCGCCGCCATCGCC 5120 5130 5140 5150 5160 5170 5180 5190 CTATCTTCAATGAAGCAATAAAGACGCCAGATTTAGATGCCATTAGCGAGGGTTCCAACGACTCTGATCTGATCGGCATC 5240 5250 5260 5270 5280 5300 5310 AGGTGAACGAAAAGGCAAATGCTTATCCCGAAGCTCCCAAACCACAACAAAAAAAA	TAAACCGACG 5120 TCTTCAATGA 5240 TGAACGAAAA	ACTAC GGC 5110 TCCAACCCTA 5230 SATCCTGAGG	CATCATCAGG 5100 TGTGGAAGGC 5220 CGTCTTCATG 5340	SO90 SO90 SACAAACACAC S210 SACTTTGGCAA S330 i	TGCCTTGTGCTACTGCTTACCTTCATCAGGACTAU 5080 5090 5100 511 GCCCCGGCACCACACACACACTGTGGAAGGCTCCAAC 5200 5210 5220 523 GAAGATCTTGGCAACGTCTTCATGGATCCT 5320 5330 1 5340 535

KEY TO UPDATED SEQUENCE:

type of change:	'G' addition	rearrangement	deletion	'C' addition	'G' addition	rearrangement	rearrangement	rearrangement	deletion	addition	addition	deletion	rearrangement	rearrangement	rearrangement	deletion	rearrangement
updated nucleotide number:	4	2105-08	2627-8	2668	2678	2946	3464	3471	4835-6	4884-4892	4938	5030	5031	5200	5216	5288	5422

Met	. Ala	Val	Asp	Val	Arg	Ile	Ala	Ala	Phe	Leu	Leu	Val	Phe	Ile	Ala
_1	,			5					10					15	
Pro	Ala	Val	Leu 20		Gln	Glu	Arg	Cys 25	Gly	Tyr	Met	Thr	Ala 30	Ile	Pro
Arg	Leu	Pro 35	Arg	Pro	Asp	Asn	Leu 40	Pro	Val	Leu	Asn	Phe	Glu	Gly	Gln
Thr	Trp 50	Ser	Gln	Arg	Pro	Leu 55		Pro	Ala	Pro	Glu 60	Arg	Asp	Asp	Leu
Cys 65	Met	Asp	Ala	Tyr	His 70	Val	Ile	Thr	Ala	Asn 75	Leu	Gly	Thr	Gln	Val 80
Ile	Tyr	Met	Asp	Glu 85	Glu	Ile	Glu	Asp	Glu 90	Ile	Thr	Ile	Ala	Ile 95	Leu
Asn	Tyr	Asn	Gly 100	Pro	Ser	Thr	Pro	Phe 105	Ile	Glu	Leu	Pro	Phe 110	Leu	Ser
Gly	Ser	Tyr 115	Asn	Leu	Leu	Met	Pro 120	Val	Ile	Arg	Arg	Val 125	Asp	Asn	Gly
Ser	Ala 130	Ser	His	His	His	Ala 135	Arg	Gln	His	Tyr	Glu 140	Leu	Pro	Gly	Met
Gln 145	Gln	Tyr	Met	Phe	Asn 150	Val	Arg	Val	Asp	Gly 155	Gln	Ser	Leu	Val	Ala 160
Gly	Val Ced		Leu	Ala 165	Ile	Val	Asn	Ile	Asp 170	Asp	Asn	Ala	Pro	Ile 175	Ile
Gln	Asn	Phe	Glu 180	Pro	Cys	Arg	Val	Pro 185	Glu	Leu	Gly	Glu	Pro 190	Gly	Leu
Thr	Glu	Cys 195	Thr	Tyr	Gln	Val	Ser 200	Asp	Ala	Asp	Gly	Arg 205	Ile	Ser	Thr
Glu	Phe 210	Met	Thr	Phe	Arg	Ile 215	Asp	Ser	Val	Arg	Gly 220	Asp	Glu	Glu	Thr
Phe 225	Tyr	Ile	Glu	Arg	Thr 230	Asn	Ile	Pro	Asn	Gln 235	Trp	Met	Trp	Leu	Asn 240
Met	Thr	Ile	Gly	Val 245	Asn	Thr	Ser	Leu	Asn 250	Phe	Val	Thr	Ser	Pro 255	Leu
His	Ile	Phe	Ser 260	Val	Thr	Ala	Leu	Asp 265	Ser	Leu	Pro	Asn	Thr 270	His	Thr
Val	Thr	Met 275	Met	Val	Gln	Val	<u>Ala</u> 280	Asn	Val	Asn	Ser	Arg 285	Pro	Pro	Arg





Trp Leu Glu Ile Phe Ala Val Gln Gln Phe Glu Glu Lys Ser Tyr Gln
290 295 300

Asn Phe Thr Val Arg Ala Ile Asp Gly Asp Thr Glu Ile Asn Met Pro 305 310 315 320

Ile Asn Tyr Arg Leu Ile Thr Asn Glu Glu Asp Thr Phe Phe Ser Ile 325 330 335

Glu Ala Leu Pro Gly Gly Lys Ser Gly Ala Val Phe Leu Val Ser Pro 340 345 350

Ile Asp Arg Asp Thr Leu Gln Arg Glu Val Phe Pro Leu Thr Ile Val

Ala Tyr Lys Tyr Asp Glu Glu Ala Phe Ser Thr Ser Thr Asn Val Val 370 380 Ced 4

Ile Ile Val Thr Asp Ile Asn Asp Gln Arg Pro Glu Pro Ile His Lys 385 390 395 400

Glu Tyr Arg Leu Ala Ile Met Glu Glu Thr Pro Leu Thr Leu Asn Phe
405 410 415

Asp Lys Glu Phe Gly Phe His Asp Lys Asp Leu Gly Gln Asn Ala Gln 420 425 430

Tyr Thr Val Arg Leu Glu Ser Val Asp Pro Pro Gly Ala Ala Glu Ala 435 440 445

Phe Tyr Ile Ala Pro Glu Val Gly Tyr Gln Arg Gln Thr Phe Ile Met
450 455 460

Gly Thr Leu Asn His Ser Met Leu Asp Tyr Glu Val Pro Glu Phe Gln 465 470 480

Ser Ile Thr Ile Arg Val Val Ala Thr Asp Asn Asn Asp Thr Arg His

Val Gly Val Ala Leu Val His Ile Asp Leu Ile Asn Trp Asn Asp Glu
500 505 510

Gln Pro Ile Phe Glu His Ala Val Gln Thr Val Thr Phe Asp Glu Thr 515 520 525

Glu Gly Glu Gly Phe Phe Val Ala Lys Ala Val Ala His Asp Arg Asp 530 540

Ile Gly Asp Val Val Glu His Thr Leu Leu Gly Asn Ala Val Asn Phe 545 550 555 560

Leu Thr Ile Asp Lys Leu Thr Gly Asp Ile Arg Val Ser Ala Asn Asp 565 570 575

Ser Phe Asn Tyr His Arg Glu Ser Glu Leu Phe Val Gln Val Arg Ala 585 Thr Asp Thr Leu Gly Glu Pro Phe His Thr Ala Thr Ser Gln Leu Val 600 Ile Arg Leu Asn Asp Ile Asn Asn Thr Pro Pro Thr Leu Arg Leu Pro 610 615 Arg Gly Ser Pro Gln Val Glu Glu Asn Val Pro Asp Gly His Val Ile 630 Thr Gln Glu Leu Arg Ala Thr Asp Pro Asp Thr Thr Ala Asp Leu Arg 650 Phe Glu Ile Asn Trp Asp Thr Ser Phe Ala Thr Lys Gln Gly Arg Gln 660 Ala Asn Pro Asp Glu Phe Arg Asn Cys Val Glu Ile Glu Thr Ile Phe 675 680 Pro Glu Ile Asn Asn Arg Gly Leu Ala Ile Gly Arg Val Val Ala Arg 695 Glu Ile Arg His Asn Val Thr Ile Asp Tyr Glu Glu Phe Glu Val Leu 710 715 Ser Leu Thr Val Arg Val Arg Asp Leu Asn Thr Val Tyr Gly Asp Asp 725 730 Tyr Asp Glu Ser Met Leu Thr Ile Thr Ile Ile Asp Met Asn Asp Asn 740 _ Ce# 7 745 Ala Pro Val Trp Val Glu Gly Thr Leu Glu Gln Asn Phe Arg Val Arg 760 Glu Met Ser Ala Gly Gly Leu Val Val Gly Ser Val Arg Ala Asp Asp 770 775 Ile Asp Gly Pro Leu Tyr Asn Gln Val Arg Tyr Thr Ile Phe Pro Arg 785 790 795 800 Glu Asp Thr Asp Lys Asp Leu Ile Met Ile Asp Phe Leu Thr Gly Gln 810 Ile Ser Val Asn Thr Ser Gly Ala Ile Asp Ala Asp Thr Pro Pro Arg Phe His Leu Tyr Tyr Thr Val Val Ala Ser Asp Arg Cys Ser Thr Glu 835 840. Asp Pro Ala Asp Cys Pro Pro Asp Pro Thr Tyr Trp Glu Thr Glu Gly 850 855

	IIe Cad		Ile	His	11e 870	Thr	Asp	Thr	Asn	Asn 875	Lys	Val	Pro	Gln	Ala 880
Glu	Thr	Thr	Lys	Phe 885	Asp	Thr	Val	Val	Tyr 890	Ile	Tyr	Glu	Asn	Ala 895	Thr
His	Leu	Asp	Glu 900	Val	Val	Thr	Leu	Ile 905	Ala	Ser	Asp	Leu	Asp 910	Arg	Asp
Glu	Ile	Tyr 915	His	Thr	Val	Ser	Tyr 920	Val	Ile	Asn	Tyr	Ala 925	Val	Asn	Pro
Arg	Leu	Met	Asn	Phe	Phe	Ser	Val	Asn	Ara	Glu	Thr	Glv	T.e.11	Va 1	Tur
J	930					935					940	<u> </u>			
Val	Asp	Tyr	Glu	Thr	Gln	Gly	Ser	Gly	Glu	Val	Leu	Asp	Arg	Asp	Glv
945					950					955			,	•	960
Asp	Glu	Pro	Thr	His 965	Arg	Ile	Phe	Phe	Asn 970	Leu	Ile	Asp	Asn	Phe 975	Met
Gly	Glu	Gly	Glu 980	Gly	Asn	Arg	Asn	Gln 985	Asn	Asp	Thr		Val 920	Leu	Val
Ile	Leu	Leu 995	Asp	Val	Asn	Asp	Asn 1000		Pro	Glu	Leu	1	Pro	Pro	Ser
Glu	Leu 1010		Trp	Thr	Ile	Ser 1015		Asn	Leu	Lys	Gln 1020	_	Val	Arg	Leu
Glu 1025		His	Ile	Phe	Ala 1030		Asp	Arg	Asp	Glu 1035		Asp	Thr	Asp	Asn 1040
Ser	Arg	Val	Gly	Tyr 1045		Ile	Leu	Asn	Leu 1050		Thr	Glu	Arg	Asp 1055	
Glu	Val	Pro	Glu 1060	Leu)	Phe	Val		Ile 1065		Ile	Ala	Asn	Val 1070		Gly
Glu	Leu	Glu 1075		Ala	Met	Asp	Leu 1080	_	Gly	Tyr	Trp	Gly 1085		Tyr	Ala
Ile	His 1090		Arg	Ala	Phe	Asp 1095		Gly	Ile	Pro	Gln 1100		Ser	Met	Asn
Glu	Thr	Tvr	Glu	Leu	Ile	Ile	His	Pro	Phe	Asn	Tvr	Tvr	Ala	Pro	Glu
1105	2_1	10			1110					1115		-1		,	1120
1	_		Pro	Thr 1125			Ala	Val	Ile 1130	-	Leu	Ala	Arg	Glu 1135	_
Ala	Val	Ile	Asn 1140	Gly	Val	Leu	Ala	Thr 1145		Asn	Gly	Glu	Phe 1150		Glu

Arg Ile Ser Ala Thr Asp Pro Asp Gly Leu His Ala Gly Val Val Thr 1155 1160 Phe Gln Val Val Gly Asp Glu Glu Ser Gln Arg Tyr Phe Gln Val Val 1175 1180 Asn Asp Gly Glu Asn Leu Gly Ser Leu Arg Leu Leu Gln Ala Val Pro 1185 1195 Glu Glu Ile Arg Glu Phe Arg Ile Thr Ile Arg Ala Thr Asp Gln Gly 1205 1210 Thr Asp Pro Gly Pro Leu Ser Thr Asp Met Thr Phe Arg Val Val Phe 1220 1225 cad 11 Val Pro Thr Gln Gly Glu Pro Arg Phe Ala Ser Ser Glu His Ala Val 1240 Ala Phe Ile Glu Lys Ser Ala Gly Met Glu Glu Ser His Gln Leu Pro 1250 1255 Leu Ala Gln Asp Ile Lys Asn His Leu Cys Glu Asp Asp Cys His Ser 1275 Ile Tyr Tyr Arg Ile Ile Asp Gly Asn Ser Glu Gly His Phe Gly Leu 1285 1290 Asp Pro Val Arg Asn Arg Leu Phe Leu Lys Lys Glu Leu Ile Arg Glu 1305 Gln Ser Ala Ser His Thr Leu Gln Val Ala Ala Ser Asn Ser Pro Asp 1315 1320 Gly Gly Ile Pro Leu Pro Ala Ser Ile Leu Thr Val Thr Val Thr Val 1335 1330 1340 Arg Glu Ala Asp Pro Arg Pro Val Phe Val Arg Glu Leu Tyr Thr Ala 1350 1355 Gly Ile Ser Thr Ala Asp Ser Ile Gly Arg Glu Leu Leu Arg Leu His 1365 1370 Ala Thr Gln Ser Glu Gly Ser Ala Ile Thr Tyr Ala Ile Asp Tyr Asp 1380 Thr Met Val Val Asp Pro Ser Leu Glu Ala Val Arg Gln Ser Ala Phe 1395 1400 1405 Val Leu Asn Ala Gln Thr Gly Val Leu Thr Leu Asn Ile Gln Pro Thr 1415 Ala Thr Met, His Gly Leu Phe Lys Phe Glu Val Thr Ala Thr Asp Thr 1425 1435 1430 1440

FIGURE 2

dc-58851

Ala Gly Ala Gln Asp Arg Thr Asp Val Thr Val Tyr Val Val Ser Ser 1445 1450 1455

Gln Asn Arg Val Tyr Phe Val Phe Val Asn Thr Leu Gln Gln Val Glu 1460 1465 1470

Asp Asn Arg Asp Phe Ile Ala Asp Thr Phe Ser Ala Gly Phe Asn Met 1475 1480 1485

Thr Cys Asn Ile Asp Gln Val Val Pro Ala Asn Asp Pro Val Thr Gly 1490 1495 1500

Val Ala Leu Glu His Ser Thr Gln Met Ala Ala Thr Ser Tyr Gly Thr 1505 1510 1515 1520

Thr Tyr Pro Tyr Ser Leu Met Arg 1525

.000			
į	:		-
		=	=
:	:	=	=
	:		
٠	:	÷	
-	;	· · · · · · · · · · · · · · · · · · ·	
	:	:	
		=	=
2	:	=	=
ŧ	:		
*******	:	3	į
******	:	:	
trees.	:	=	
Ŀ		=	ì
•		=	

MPPIFVPENGK		YVTOPPRATSALLRVTVLDTN YVTOP	GLDPVRNRLFLKKELIRKDSASHTLQVAASNSPDGGIPLPASILTVTVTVREADPREM Sus Motif GDREDREDRE
fat EC18 fat EC18 DC42 EC2 HPT-2 EC2 BTRCad-1 BTRCad-2 BTRCad-4 BTRCad-6 BTRCad-6 BTRCad-6 BTRCad-6 BTRCad-7 BTRCad-9 BTRCad-10	P EC1	1087651 1087651 1087651	TRcad- adhèri

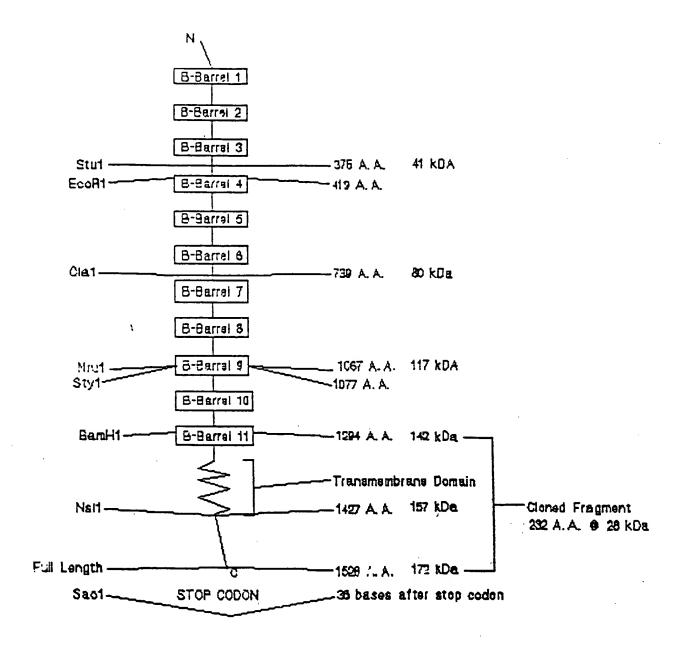


Fig. 3 Block diagram of cadherin-like structure of BT-R₁

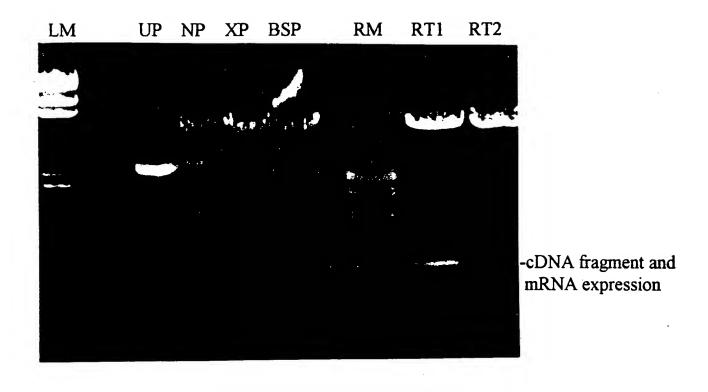


Fig. 7 Clone characterization of BamHI-SacI fragment of BT-R₁. LM is HindIII cut Lambda marker; UP is the uncut plasmid clone; NP is NsiI cut plasmid; XP is XhoI cut plasmid; BSP is BamHI and SacI cut plasmid showing the cloned fragment from BT-R₁; RM is mRNA size marker; and RT1 and RT2 are transcribed mRNAs from the cloned BT-R₁ fragment.

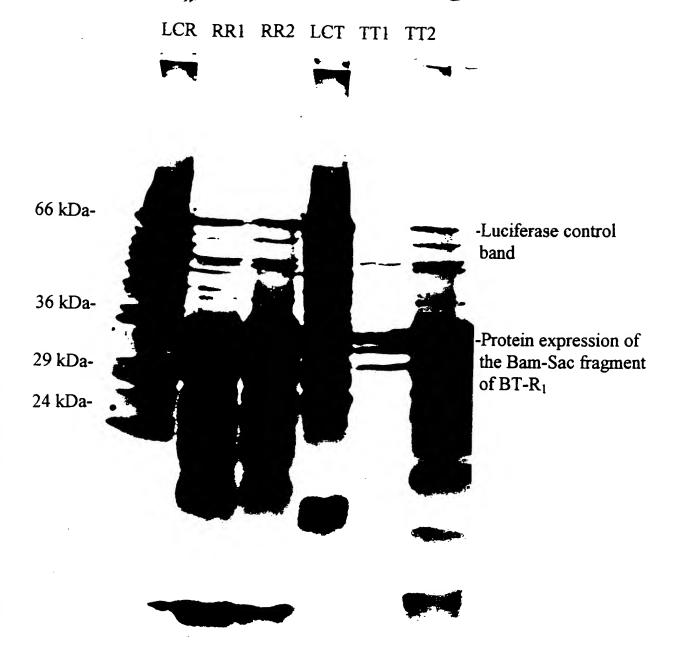


Fig. 5 Detection of protein expression from the plasmid containing the Bam-Sac fragment of BT-R₁ using ³⁵S-methionine as a tag. LCR is a luciferase control mRNA to show that the rabbit reticulocyte lysates are functional; RR1 and RR2 are expression products of the Bam-Sac fragment of BT-R₁ produced in rabbit reticulocytes from mRNA; LCT is a luciferase control plasmid to show that the transcription/translation kit is functional; and TT1 and TT2 are expression products of the Bam-Sac fragment of BT-R₁ produced in a transcription/translation kit.

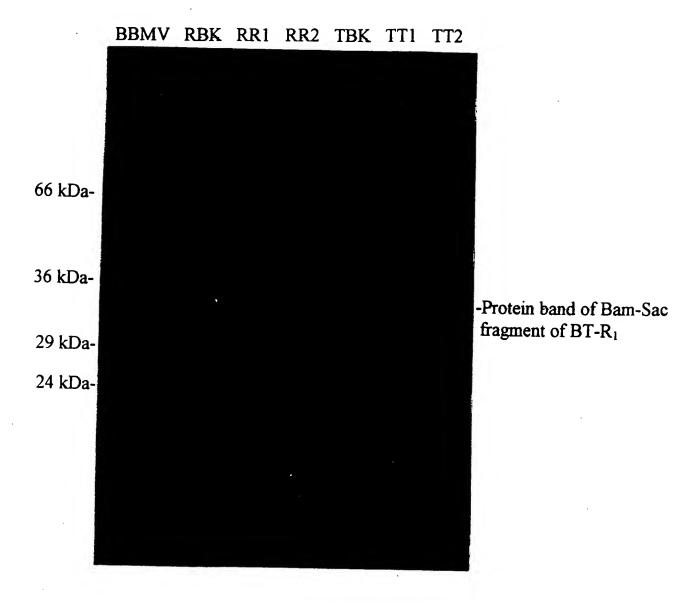
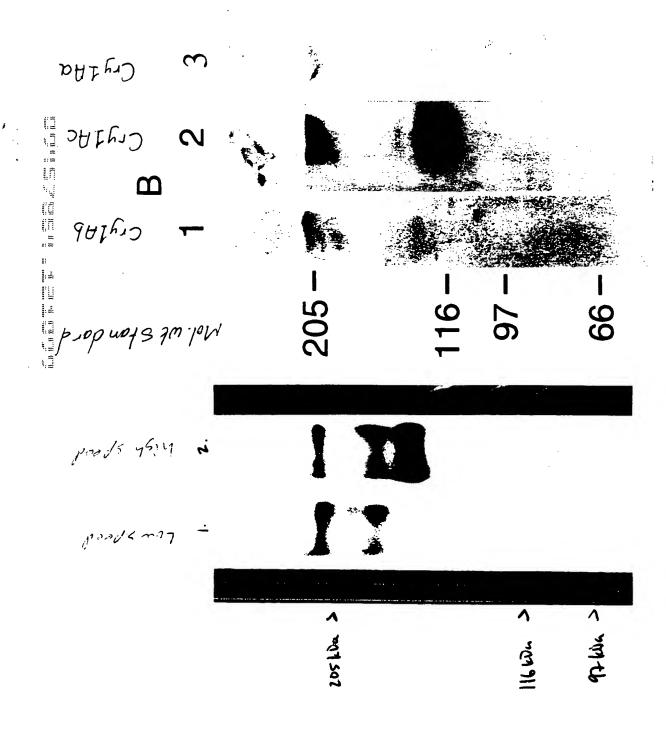


Fig. 6 Radio-blot of the Bam-Sac fragment of BT-R₁ with ¹²⁵I-labeled Cryl Ab. BBMV is the brush border membrane vesicles from the midgut of M. Sexta containing the wild-type BT-R₁ receptor protein; RBK is a rabbit reticulocyte blank; RR1 and RR2 are the expression products of the Bam-Sac fragment of BT-R₁ produced in rabbit reticulocytes from mRNA; TBK is a transcription/translation kit blank; TT1 and TT2 are expression products of the Bam-Sac fragment of BT-R₁ produced in a transcription/translation kit. The arrows point to two of the bands.

(un borer



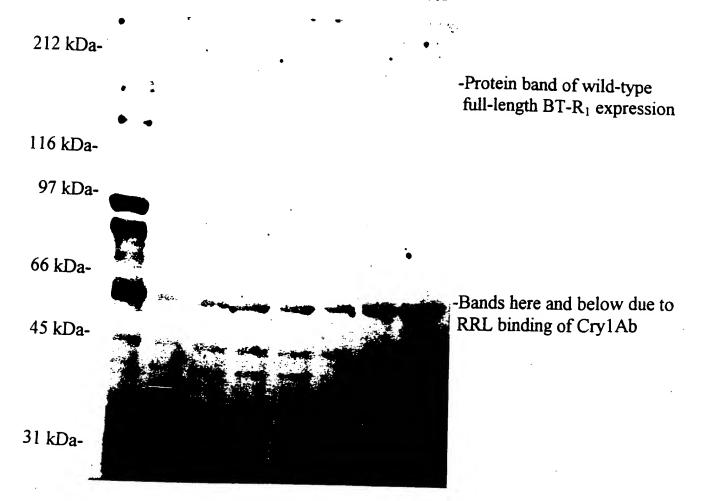
Pink bollworm (G C Pectunophora gossy [la) were Separated by s 7.5% SDS-PHGE, blothed and probed with 2x105 cpm per m1 of "25 I. labeled Cry Taxino.

BBMV praterno

(2004g) from the

E. H. H. H. H. H. H. H.

WT NT BT NrT CT XT ST BLK



Radio-blot of BT-R₁ and truncation mutants of BT-R₁ with ¹²⁵I-labeled CrylAb. WT is the wild-type full-length BT-R₁ receptor; NT is the truncation mutant resulting from NsiI digestion; BT is the mutant made with BamHI; NrT is the mutant made with NruI; CT is the mutant made with ClaI; XT is the mutant made with XhoI; ST is the mutant made with StuI; and, BLK is rabbit reticulocyte lysates containing only endogenous proteins.